



NCAR

Data QC & Archiving: Upcoming Developments

Seth McGinnis

3rd NARCCAP Users Workshop

NCAR

Boulder, CO

April 7-8, 2011

National Center for Atmospheric Research

The Data Pipeline

- Simulation
- Data transfer
- QC
- Publishing



QC

- Tally variables & files
- cf-checker
- Time coordinate continuity
- Temporal min/max/mean
- Time-space cross-sections



Fixing errors

- **Metadata errors:** fix
- **Data gaps:** fill with `missing_value`
- **Bad values:** replace with `missing_value`
- **More serious errors:** contact modelers

Currently in the Pipeline

- CAM timeslice: current run table 5 (3D)
- CRCM: ccsm/cgcm3 – table 3,5
- ECP2: gfdl current – tas, pr
- HRM3: hadcm3 – everything not yet pub.
- MM5I: ccsm – radiation fluxes
- RCM3: gfdl-rerun/cgcm3 – table 2,3 (2D)
- WRFG: cgcm3 – table 5 (3D)

Derived Data Products

- Derived variables (e.g., wind speed)
- Temporal averages & climatologies
- Regridded to 0.5-degree lat/lon grid
- Shapefiles on NCAR GIS portal

Other Stuff

- Data Citation
- DOI for the dataset
- Sharing tools, processes